Analyzing Hand Color for Health Status Prediction: A Non-Invasive ML Approach

- Capstone 3 Project
- Adli Karadsheh
- August 02, 2025

Problem Statement & Motivation

- Healthcare challenge: Early detection of circulatory/oxygenation issues (e.g., hypoxia) via hand color—bluish (unhealthy) vs. pink (healthy).
- Real-world impact: Non-invasive monitoring for at-risk patients (cardiovascular/respiratory conditions); potential mobile app for remote screening.
- DSM Alignment: Problem defined, stakeholders (patients/providers), goals (>80% accuracy), risks (imbalance/lighting).

Dataset Overview

- Sources: ~70 healthy (stock sites: Pexels/Unsplash/Adobe); ~57 unhealthy (clinic images with bluish tones).
- Processing: MediaPipe for landmarks; RGB at midpoints (thumb: 2 segments; others: 3); radii 1/3/5 → ~84 features/CSV.
- Stats: Healthy brighter (e.g., Avg_G ~160 vs. ~140, $p < 0.001$ via t-tests); mild imbalance (~58% unhealthy).
- Wrangling: Imputed NaNs (means); no rows lost.

Radius (pixels)	Channel	t-statistic	p-value
1	Avg_R	4.22	4.72E-05
1	Avg_G	5.9	3.18E-08
1	Avg_B	4.99	1.97E-06
3	Avg_R	4.15	6.10E-05
3	Avg_G	5.89	3.35E-08
3	Avg_B	4.96	2.29E-06
5	Avg_R	4.11	7.13E-05
5	Avg_G	5.99	2.07E-08
5	Avg_B	5.06	1.45E-06

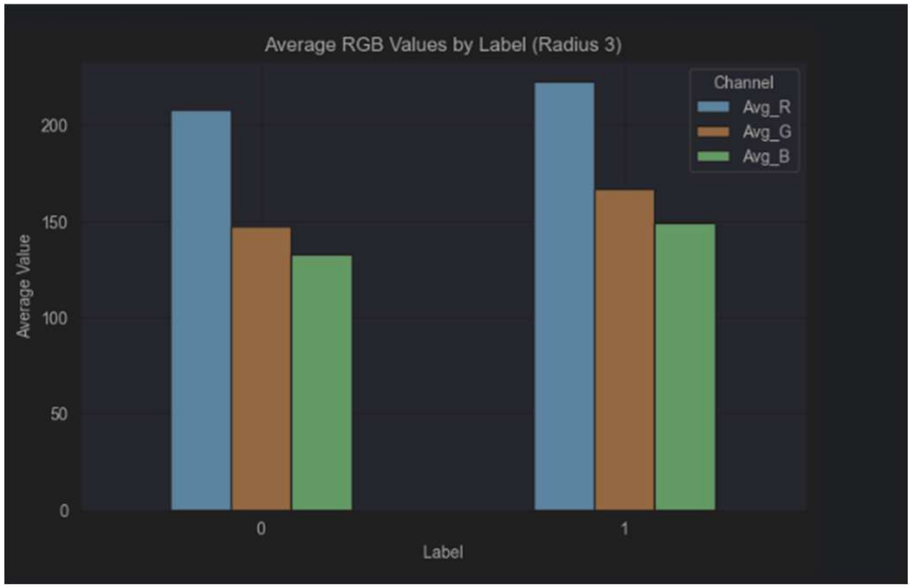
Exploratory Data Analysis (EDA)

•Distributions: RGB ~50-255; bimodal histograms reflecting labels.

•Correlations: High within channels (~0.9+); Label ~0.4-0.5 with red/green.

•Trends: Larger radii smooth noise (std drop ~10%); better separation (green widest gap).

•Insights: Outliers minimal; pairplots show R-G clusters.



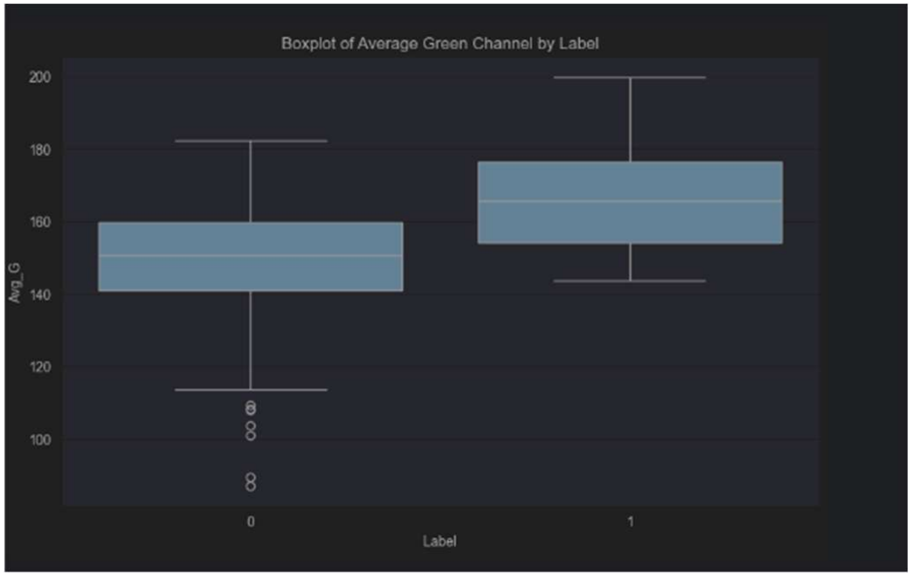
Exploratory Data Analysis (EDA)

•Distributions: RGB ~50-255; bimodal histograms reflecting labels.

•Correlations: High within channels (~0.9+); Label ~0.4-0.5 with red/green.

•Trends: Larger radii smooth noise (std drop ~10%); better separation (green widest gap).

•Insights: Outliers minimal; pairplots show R-G clusters.



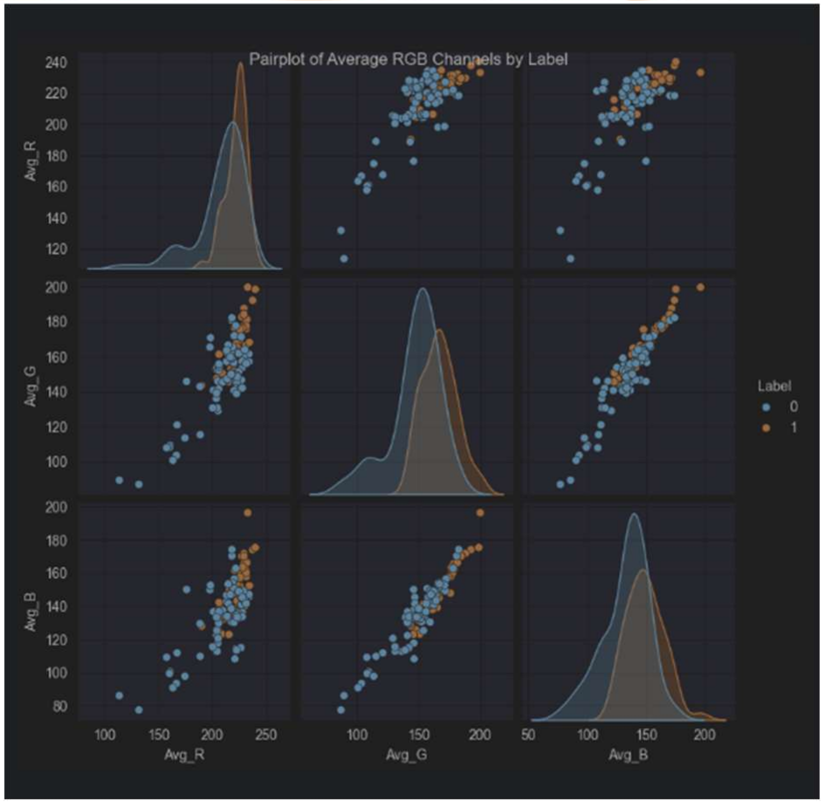
Exploratory Data Analysis (EDA)

•Distributions: RGB ~50-255; bimodal histograms reflecting labels.

•Correlations: High within channels (~0.9+); Label ~0.4-0.5 with red/green.

•Trends: Larger radii smooth noise (std drop ~10%); better separation (green widest gap).

•Insights: Outliers minimal; pairplots show R-G clusters.



Methods & Modeling Pipeline

•Preprocessing: Standardized; PCA (20 components, ~97% variance); 80/20 train-test split (stratified)
.
•Models: Initial LR, Regularized LR, Random Forest (RF), Neural Network (NN)—focus on RF as best.

•Evaluation: Test accuracy, 5-fold CV (± std), confusion matrices (low FN priority).

•Tools: Python (scikit-learn, TensorFlow, MediaPipe); notebooks for EDA/modeling.

Model Results & Comparison

- RF (Best): 92% test across radii; CV up to 83% at R5 (+11% from R1 via smoothing).
- Comparison: RF > NN/Reg LR (by +4-11% CV); importances: Comp 1 (global shifts), Comp 16 (green contrasts).
- Predictions vs. Actual: RF near-perfect (1 FN at R5); aligns with bright healthy tones.
- Limitations: Small data (variability in NN); unaccounted lighting/skin tones.

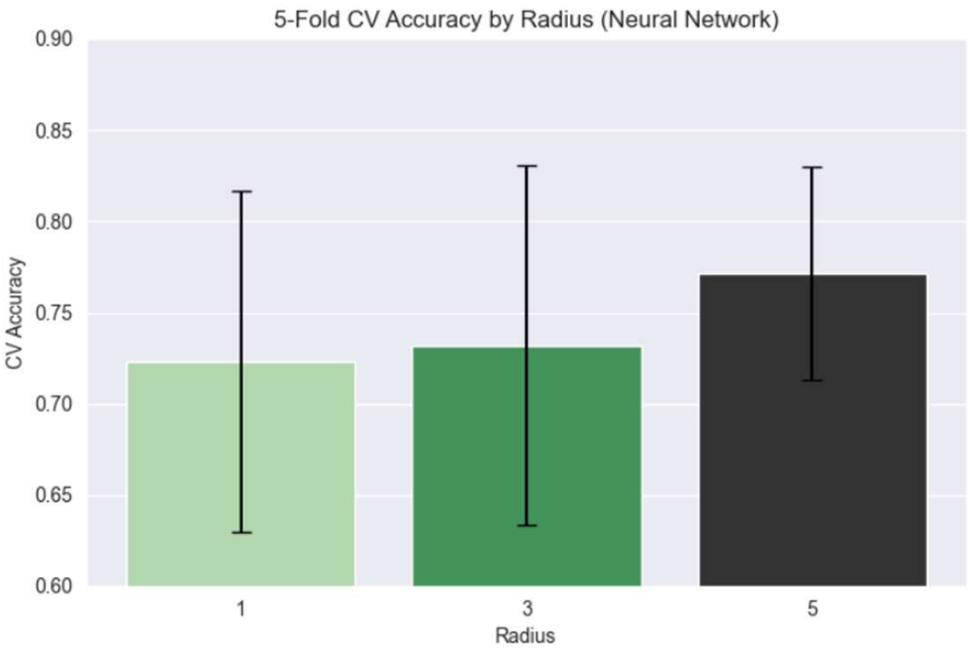
Model	Radius (pixel)	Test Accuracy	5-Fold CV Accuracy (std)	
Initial LR	1	0.85	0.73 (±0.12)	Baseline; balanced CM. Vs. RF: -7% test, linear limits.
Initial LR	3	0.88	0.74 (±0.11)	Strong initial; good TN. Vs. RF: Lower CV, no ensemble robustness.
Initial LR	5	0.85	0.72 (±0.10)	Stable; minor radius gain. Vs. RF: -7% test, misses non-linearity.
Regularized LR	1	0.85	0.73 (±0.12)	No warnings; interpretable coeffs (Comp 10 ~0.80). Vs. RF: -7% test; simpler but less accurate.
Regularized LR	3	0.88	0.74 (±0.11)	Best LR; 0 FN at test. Vs. RF: Equivalent test but -4% CV; RF's importances more insightful (Comp 16 ~0.09).
Regularized LR	5	0.85	0.72 (±0.10)	Even errors; scalable. Vs. RF: -7% test; RF better on TN (14/15 vs. 13/15).
Random Forest (Current)	1	0.92	0.72 (±0.10)	Vs. Reg LR: +7% test; Comp 1/10 key. Vs. NN: Less variance, better TN.
Random Forest (Current)	3	0.92	0.76 (±0.11)	Vs. Reg LR: +4% CV; importances shift to Comp 16. Vs. NN: +7% test, consistent CM.
Random Forest (Current)	5	0.92	0.83 (±0.12)	Top: +11% CV gain; Comp 16/8-9 dominant. Vs. All: Highest metrics, ideal for deployment—low errors, robust to radius.
Neural Network	1	0.88	~0.72 (±~0.10)	Non-linear; stochastic. Vs. RF: -4% test, more variance.
Neural Network	3	0.81	~0.70 (±~0.11)	More FN. Vs. RF: -11% test; needs tuning.
Neural Network	5	0.88	~0.76 (±~0.12)	Aligns on trend. Vs. RF: -4% test; RF more consistent.

Discussion & Limitations

•Success: RGB midpoints effective (green key differentiator); RF robust for screening.

•Limitations: Dataset size risks overfit; multicollinearity; no HSV/lighting normalization.

•Variability: NN stochastic (±0.06-0.12 CV); errors in borderlines (e.g., shadows).



Next Steps & Recommendations

•**Processing**: Add HSV (for lighting invariance); resize uniformly (512x512); equalize histograms.

•**Labeling**: Include age/chronic conditions (e.g., diabetes) for multi-class; collect oximetry data.

•**Models**: Tune RF (n_estimators=200); ensemble RF+NN; add CNNs for raw images.

•**App**: Prototype in Flutter (TF Lite integration) for real-time predictions/alerts.

Conclusion

•Achieved >80% goal (RF 92%); validates color-based monitoring.

•Impact: Bridges DS/healthcare; future app could empower users.

•Thanks & Q&A.

The Hands Tell More Than Sign